干旱区地理
ARID LAND GEOGRAPHY

doi:10.12118/j.issn.1000-6060.2019.05.19

基于 CART 决策树的沙地信息提取方法研究①

张晞伟[1]， 王 磊[2]， 汪西原[1,3]
（1 宁夏大学物理与电子电气工程学院，宁夏 银川 750021； 2 西北土地退化与生态恢复省部共建国家重点实验室培育基地，宁夏 银川 750021； 3 宁夏回族自治区沙漠信息智能感知重点实验室，宁夏 银川 750021）

摘 要：为研究沙地信息提取的方法，采用基于 CART 决策树的面向对象方法，提取中卫市沙坡头区的沙地信息。首先对研究区进行多尺度分割和光谱差异分割得到对象层，然后选择合适的提取特征和训练样本点，最后输入选择的提取特征和样本点生成 CART 规则树，并对地物进行分类，提取出沙地信息。结果表明：采用面向对象的 CART 决策树方法提取沙地信息具有较高自动化程度和精确度，依此构建的 CART 决策树总体分类精度可达到77%，是最近邻分类结果的1.12倍，支持向量机分类结果的1.57倍，此外，*NDBI*（归一化裸露指数）、*GSI*（粒度指数）和 *SWIR* 2（第七波段）均值可以成功的将沙地、戈壁和裸岩石砾地三个易混地物区分开来，是沙地提取过程中三个重要的特征指数。

关 键 词： 面向对象；多尺度；光谱差异；CART 决策树；沙地提取

沙漠化是备受国际关注的重大环境问题之一，它直接影响人类的生存环境和社会的稳定。宁夏是我国风沙危害严重的省区之一，其东、西、北三面分别被毛乌素、腾格里、乌兰布和 3 大沙漠、沙地包围，范围涉及 13 个县（市、区）、90 个乡镇及国有农林场[1]。准确快捷地获取沙漠化信息，是正确制定沙漠化防治政策的重要依据，也是沙漠化研究的关键问题之一。近些年遥感监测凭借其自动化程度高、解译范围广、周期性短等优势，广泛的被应用于各种地物的提取，其中被用于沙地信息提取的遥感解译方法也在不断的推陈出新。李元科[2]通过目视解译研究预估了盐池县沙漠化动态；李晓琴[3]利用光谱混合分析技术，探讨了光谱角分类法（SAM）自动提取沙化土地和盐渍化土地信息；乔平林[4]用 BP 神经网络模型提取了内蒙内克什克腾旗西北部的沙地信息。在基于象元出现“椒盐现象”后，研究人员进而开辟了面向对象的领域[5-6]。

CART 决策树是在 1984 年由 Breman 提出的一种二叉树分类方法，它能够充分利用空间辅助信息，对影像的多个特征变量进行选择性组合，识别繁多而有效的信息，从而实现对未知数据的分类、预测和挖掘。近些年来，很多国内外专家学者做了 CART 决策树分类应用方面的研究，通过与传统分类方法的对比，得到了 CART 决策树分类方法分类精度高，速度快等结论。孙建伟[7]利用 CART 决策树方法对湖北荆州市八岭山镇进行了地物分类，取得了良好的效果，段英杰[8]、康文平[9]用 CART 决策树分别对西藏部分区域和内蒙古部分区域做了沙漠信息提取方面的研究。吕利利[10]以民勤县为例，利用 CART 决策树结合多个特征提取了沙地信息。这些沙地信息提取研究或是直接对像素层直接处理，或是提取特征较少，或是提取过程复杂自动化程度低。本文以宁夏中卫市沙坡头区为研究区，利用 Landsat8 数据源，结合面向对象的分类方法，输入多个提取特征，研究一种基于 CART 决策树高自动化提取沙地信息的方法。

1 研究区概况与数据源

1.1 研究区概况

中卫市沙坡头区位于宁夏回族自治区中西部，

① **收稿日期**：2019-02-10； **修订日期**：2019-05-24
基金项目：国家自然基金（41561087）
作者简介：张晞伟（1993-），男，硕士，研究方向为遥感图像处理研究. E-mail：zhang_xiwei@163.com
通讯作者：汪西原. E-mail：wangxiy@nxu.edu.cn

东邻中宁县,南与同心县、海原县及甘肃省靖远县交汇,西接甘肃省景泰县,北邻内蒙古自治区阿拉善左旗。境域东西长 115.3 km,南北宽 81.4 km,总面积 5 922.4 km^2。地跨 104°17′~106°10′E、36°06′~37°50′N。境内东接腾格里沙漠南侧边缘,中部有黄河截腰穿过,南边以山台、山地为主,整体地势呈西南高、东北低的势态。沙坡头区属于半干旱气候,是典型的大陆性季风气候和沙漠气候,春暖迟、秋凉早、夏热短、冬寒长,风大沙多,干旱少雨,而又得黄河灌溉之优势,成为了西北地区重要的商品粮、水产品和设施蔬菜生产基地。所以,中卫人民自古以来凭借自己的智慧和努力与沙漠化作斗争,研究沙坡头区的沙地信息对于治沙防沙具有重大意义。

1.2 数据源

Landsat-8 卫星是 2013 年 2 月 11 号由 NASA 发射的陆地卫星,Landsat-8 上携带有两个主要载荷:OLI(Operational Land Imager,陆地成像仪)和 TIRS(Thermal Infrared Sensor,热红外传感器)。OLI 陆地成像仪包括 9 个波段,空间分辨率为 30 m,其中包括一个 15 m 的全色波段,TIRS 包括 2 个波段。Landsat 8 陆地卫星凭借其客观、大尺度、多时相、低成本、方便快捷及非破坏性等特点,极其适合在土地沙漠化监测中应用。本文选取的数据是 2016 年 8 月 9 日美国陆地卫星 Landsat-8 所获取的遥感数据影像,名称代码为 LC81300342016222LGN00。在信息提取前,以 ENVI 作为图像预处理软件,对获取的影像按序进行辐射定标、FLAASH 大气校正、影像融合及裁剪。最终得到空间分辨率为 15 m 的宁夏中卫市沙坡头区多光谱数据影像(图 1)。

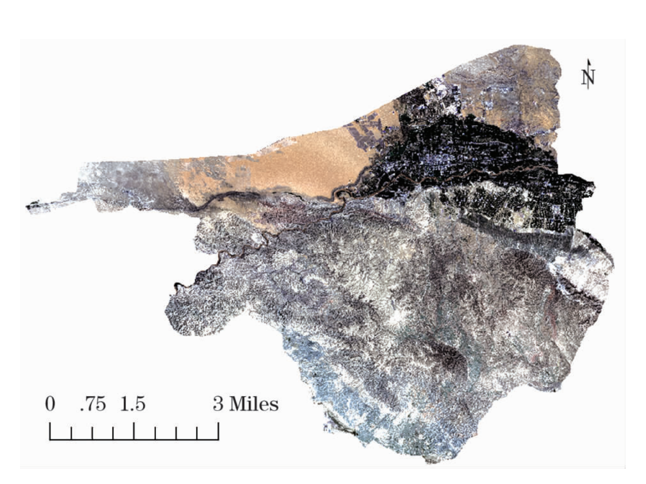


图 1 宁夏中卫市沙坡头区预处理数据

Fig. 1 Preprocessing data of Shapotou District in Zhongwei, Ningxia

2 实验研究流程及分类原理

基于 CART 决策树算法的面向对象提取方法主要包括:影像分割、对象特征提取、CART 决策树分类规则的建立、利用生成的决策树进行分类。工作采用 ENVI 和 Ecognition 软件进行。其整体技术流程如图 2 所示。

2.1 影像分割原理

2.1.1 *多尺度分割* 多尺度分割是面向对象的提取方法中较为常用的一种分割算法,它是一种自下而上(bottom-up)的方法,通过合并相邻的像素或小的分割对象,在保证对象与对象之间平均异质性最小、对象内部像元之间同质性最大的前提下,基于区域合并技术实现影像分割。异质性的具体计算方式如下:

$$f = w_{color} h_{color} + w_{shape} h_{shape} \tag{1}$$

$$h_{shape} = w_{compact} h_{compact} + w_{smooth} h_{smooth} \tag{2}$$

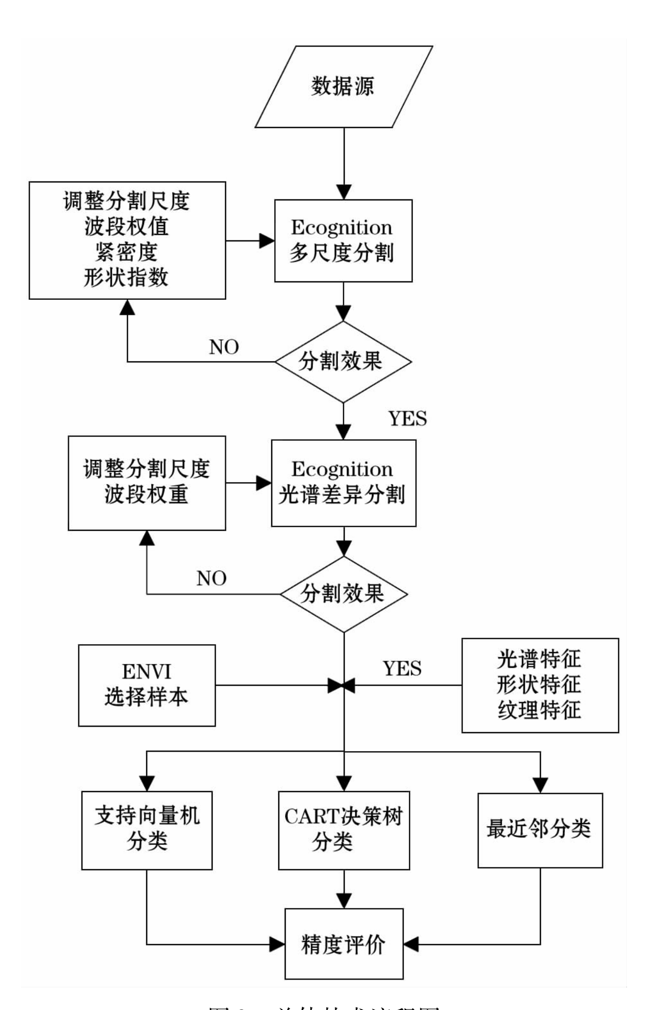


图 2 总体技术流程图

Fig. 2 General technical flow chart

式中:f 是异质性;w_{color} 是光谱信息权重;w_{shape} 是形状信息权重;这两个权重的和为 1;h_{color} 是光谱异质性的值;h_{shape} 是形状异质性的值;$w_{compact}$ 和 w_{smooth} 分别是紧密度和光滑度的权重,它们的和为 1;$h_{compact}$ 和 h_{smooth} 分别是紧密度和光滑度的异质性值[11]。

2.1.2 光谱差异分割 光谱差异分割不能基于像素层来创建新的分割层,它仅仅是一种分割优化手段,它需要在多尺度分割算法得到的分割结果基础上,通过分析相邻分割对象的亮度差异是否满足给定的阈值,来决定是否将对象进行合并。相邻影像对象之间的光谱差异(g)的计算公式为:

$$g = wn_{b1} \cdot |b1_1 - b1_2| + wn_{b2} \cdot |b2_1 - b2_2| + wn_{b3} \cdot |b3_1 - b3_2| \quad (3)$$

式中:$b1$、$b2$、$b3$ 分别表示影像的 3 个波段;wn_{b1}、wn_{b2}、wn_{b3} 分别是三个波段的归一化后的权重。归一化计算如下:

$$wn_{b1} = w_{b1}/(w_{b1} + w_{b2} + w_{b3}) \quad (4)$$

$$wn_{b2} = w_{b2}/(w_{b1} + w_{b2} + w_{b3}) \quad (5)$$

$$wn_{b3} = w_{b3}/(w_{b1} + w_{b2} + w_{b3}) \quad (6)$$

利用光谱差异分割,一方面能够减少分割对象的数量,另一方面也能在一定程度上改善"过分割"的现象[12]。

2.2 CART 决策树分类原理

分割只是将像素层划分为一个一个的对象,如果将其赋予类别信息,还需要分类这一关键环节。本文采用 CART 决策树方法对影像对象进行分类。CART(Classification And Regression Tree,分类回归树)算法采用了一种二分递归分割的技术,通过计算 Gini 系数选取属性集中的某个属性,将当前待分类的样本集分为两个子样本集,并循环此步骤,直到当前待分类的样本集被判定为叶节点或达到停止分类的条件。

基尼系数的定义为:

$$Gini_Index = 1 - \sum_{j}^{J} P^2(j/h) \quad (7)$$

$$P(j/h) = \frac{n_j(h)}{n(h)},\ \sum_{j}^{J} P(j/h) = 1 \quad (8)$$

式中:$P(j/h)$ 是从训练样本集中随机抽取一个样本,当某一测试变量值为 h 时属于第 J 类的概率;$n_j(h)$ 为训练样本中测试变量值为 h 时属于第 J 类的样本个数;$n(h)$ 为训练样本中该测试变量值为 h 的样本个数;J 为类别个数[13]。

由于每次分割都只将样本集划分为两个子集,因此最终生成的决策树是一个二叉树,结构简洁明了,使用的分割规则清晰易懂,整个过程也比较符合人们一般的思维习惯。

2.2.1 训练样本选择 CART 决策树分类是一种典型的监督分类算法,训练样本选择的好坏直接影响分类的优劣,所以样本的选择不仅要具有代表性和典型性,而且还要具有区域样点的完备性[14]。通过对研究区域的地表覆盖状况进行解译,将研究区分为沙地、植被、水域、建设用地、戈壁和裸岩石砾地 6 类土地类型,在此分类体系下,本文利用 Google 地图,用人机交互的方法选择 1 552 样本点作为训练对象,其中,沙地、植被、水域、建设用地、戈壁和裸岩石砾地的样本数分别为:158、248、42、321、345、438 个。

研究区地表覆盖类型复杂多样,沙地的样本来自流动沙地;戈壁包括细石沙砾和没有农作物的耕地;植被以草地、林地、绿色耕地为主;水域是湖泊、黄河、水库等;建设用地有城镇、道路和其他设施用地;裸岩石砾地为山丘、山脉。所选样本均匀分布在研究区内,并且代表每一类别的特点区域。此外,这些样本点也将会作为对比算法(最近邻法和支持向量机法)的样本,从而达到使精度评价具有客观性的目的。

2.2.2 特征指标 目前尚没有能够直接准确提取沙地信息指数的方法,一般运用分类方法提取。CART 决策树使用贪心算法,从大量属性中选择最优组合构成分类模型,数据特征变量的增加可能会导致决策规则的改变,但最终得到的分类方法一定是最优。因此,在遥感影像的分类中,尽可能提供有效特征量对增加决策树精度具有很大的帮助。总结相关研究,本文从光谱特征、形状特征、纹理特征、自定义特征这几个方面选取,总共 22 个特征。

光谱特征:原影像的 7 个波段均值。

形状特征:选取了紧致度(compactness)和形状指数(shape index)。

纹理特征:采用了灰度共生矩阵(GLCM)的几个指标,包括均值(Mean)、熵(Entropy)、同质性(Homogeneity)、相异性(Dissimilarity)、方差(Std-Dev)、对比度(Contrast)、相关性(Correlation)。

自定义特征：用于提取植被的归一化植被指数（*NDVI*）和土壤调节植被指数（*SAVI*）[15]；常用于提取水体的修改型归一化水体指数（*MNDWI*）[16]；用于反映地表的裸露状况的归一化裸露指数（*NDBI*）和粒度指数（*GSI*）[17-18]；用于提取城镇的比值居民地指数（*RRI*）[19]。

3　分类实现及结果分析

3.1　多尺度分割及光谱差异分割

本试验采用试错法选取最优分割尺度参数，即通过不断变换分割的尺度以及调整形状和紧密度的权重，得到不同的分割结果，最后根据先验知识和人的判别能力选取最优分割尺度参数[20]。以 5 为步长，从尺度 5 取到尺度 100，经过多次试验发现，在形状信息权重为 0.2，紧密度为 0.5 的情况下，尺度为 70 时会出现戈壁与其周边沙地混为一个对象的现象，尺度为 30 时地物分割过于破碎，50 的尺度下，同一个对象满足异质性比较小的条件，而又有些过分割，为下一步光谱差异分割做好准备，如图 3 所示。

影像分割中各个波段的权重选择对分割结果也会产生很大影响，由于本文选用分类的方法提取沙地信息，所以对每个波段都有同样的诉求，因此设置七个波段的比值为 1∶1∶1∶1∶1∶1∶1[5]。然后对尺度为 50 的结果进行光谱差异分割，同样采用试错法，尽量选择可以将若干个相同地物类型的对象合并为一个，或少数对象，错分概率小的尺度，如图 4。最终选择尺度为 100 的优化分割结果作为后续分类的目标图像。

3.2　影像分类

把在 ENVI 里选好的样本导入到 Ecognition 里，Ecognition 提供自动化程度很高的决策树分类器，选择上文选取的 22 个特征指数，做 CART 决策树分类，并生成分类规则。分类规则如图 5，决策树左枝为是，右支为否。

观察生成的规则树得出：（1）通过 *NIR*（第五波段）和 *SWIR* 1（第六波段）均值可区分出水域和裸岩石砾地；（2）*NDBI*、*RRI*、*NIR*（第五波段）均值和 *RED*（第四波段）均值的组合提取出植被和建设用地；（3）由 *GSI* 和 *SWIR* 2（第七波段）均值分类出沙地和戈壁，由 *RRI* 的另一阈值、*Coastal*（第一波段）均值、*SWIR* 2（第七波段）均值、*NDBI*、*GSI* 分离出裸岩石砾地、建设用地、戈壁。其中，容易混淆的三类地物：沙地、戈壁和裸岩石砾地，主要由 *NDBI*、*GSI* 和 *SWIR* 2（第七波段）均值三个特征所区分。基于以上规则树，生成分类结果，如图 6a 所示。

为比较 CART 决策树的分类精度，并确保评价的客观性和精确性，本文基于相同的分割图层和相同的样本，选用同样的特征，分别采用最近邻法和支持向量机法两种经典的监督分类方法，对研究区做沙地提取。其中最近邻分类方法中进行特征空间最优化（Feature space optimization），筛选出 5 个特征（Compactness、Shape index、Mean Layer5、Mean Layer1、*NDVI*）并应用于所有类别，其分类结果如图 6b 所示。支持向量机分类方法采用与 CART 决策树同样的特征，分类结果如图 6c 所示。

3.3　精度评价

观察图 6，CART 决策树和最近邻分类效果较为满意，而支持向量机的分类结果中，除了沙地和水域提取效果不错外，其他都不是很好，尤其黄河以南建设用地和戈壁还有裸岩石砾地混淆比较严重。从定量方面研究分析，结合目视解译随机选择 554 个检验点，为保证精度评价的准确性，检验点的选择应尽量选择纯净像元。本文采用混淆矩阵对 3 种分类结果分别进行精度评价，汇总 3 种分类结果的精度评价，如表 1 所示，利用 CART 决策树分类 104 个沙地检验点无 1 个出错，其生产者精度达到了 100%，用户精度达到了 98%，总体分类精度为 77%，Kappa 系数为 0.72。而利用最近邻分类 104 个沙地检验点有 35 个错分，其生产精度为 66%，用户精度为 98%，总体分类精度为 69%，Kappa 系数为 0.63。对于支持向量机分类方法，104 个沙地检验点也无一出错，用户精度为 86%，其总体分类精度为 49%，Kappa 系数为 0.38。由此可见，CART 决策树分类方法是一种高效高精度的沙地提取方法。

4　结 论

本文通过对 CART 决策树研究和讨论，采用面向对象方法，成功构建了 CART 决策树模型，并将其应用于中卫市沙坡头区的沙地提取。在此得出以下结论：

（1）对比最近邻分类和支持向量机分类，CART 决策树表现出了非常好的分类精度，总体分类精度分别是最近邻和支持向量机的 1.12 倍、1.57 倍，Kappa 系数分别是它们的 1.14 倍、1.89 倍。尤其在

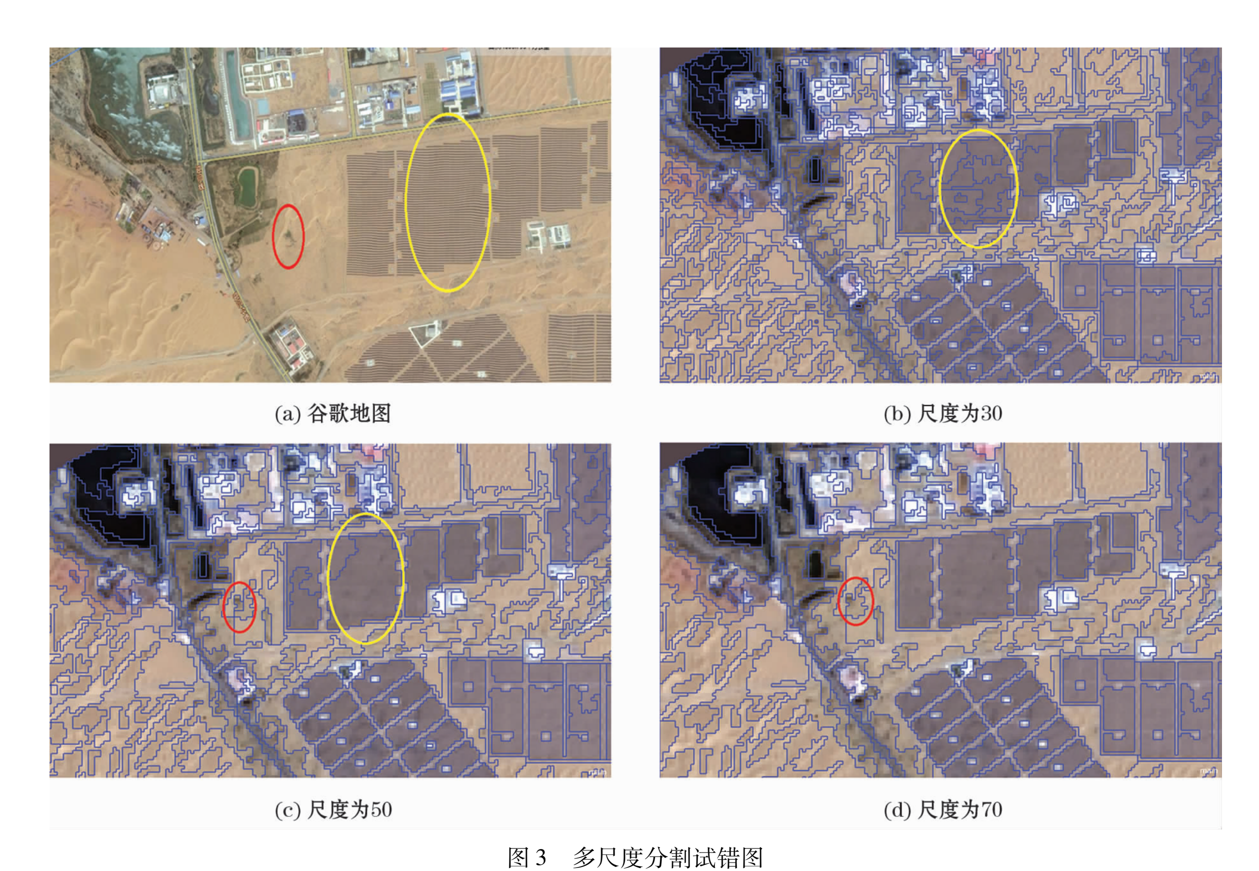


图 3　多尺度分割试错图

Fig. 3　Multi-scale segmentation trial and error chart

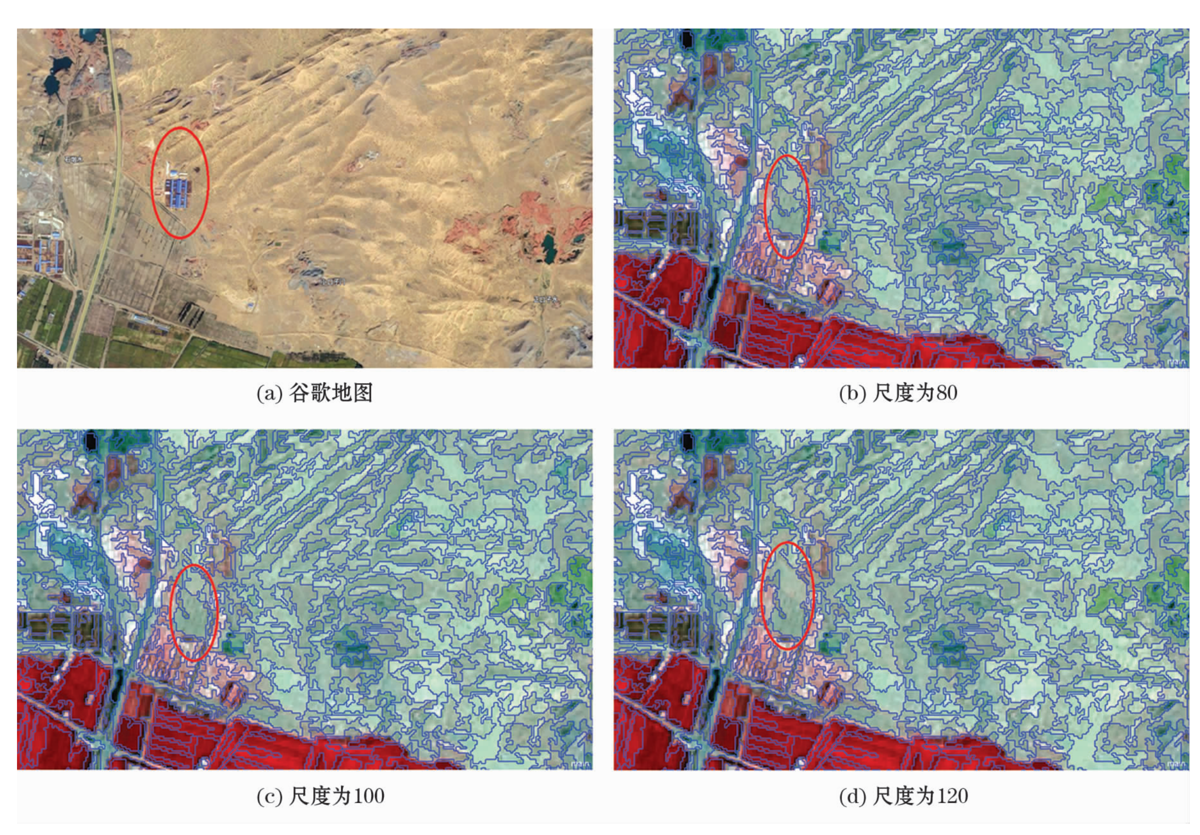


图 4　光谱差异分割试错图

Fig. 4　Spectral difference segmentation trial and error chart

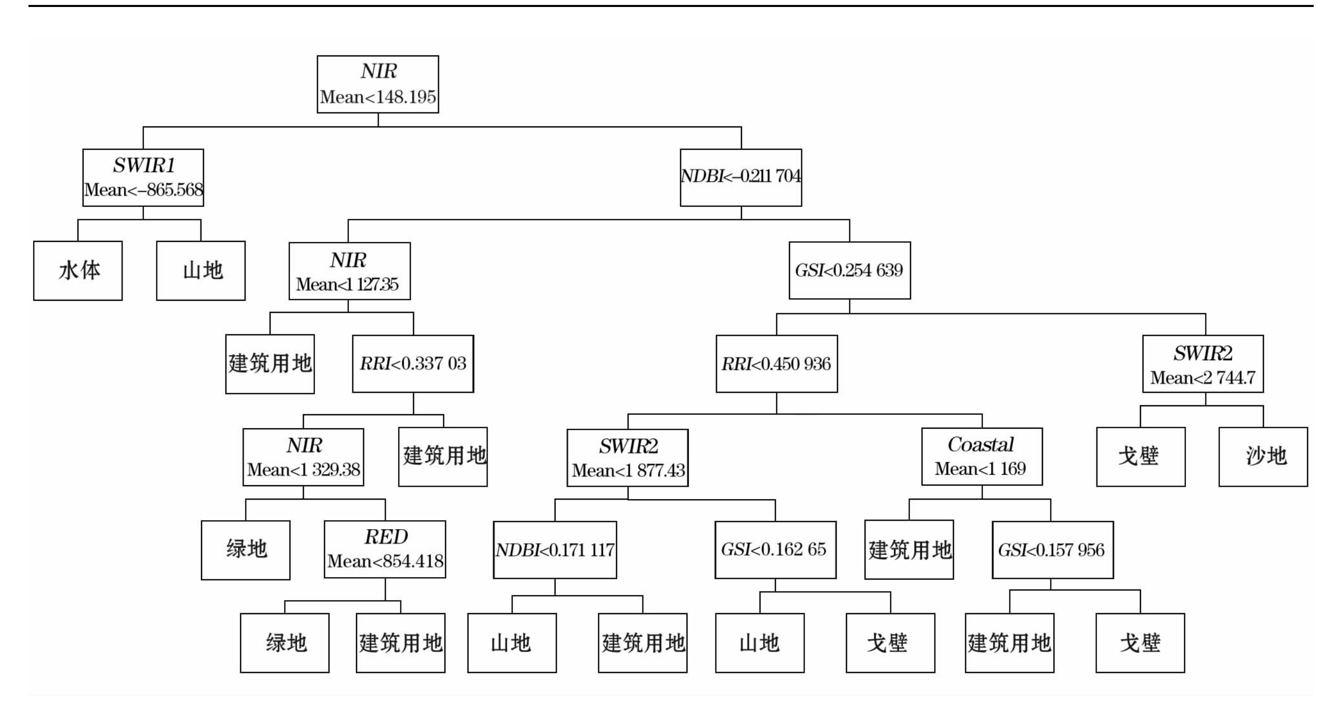


图 5　CART 规则树

Fig. 5　CART rule tree

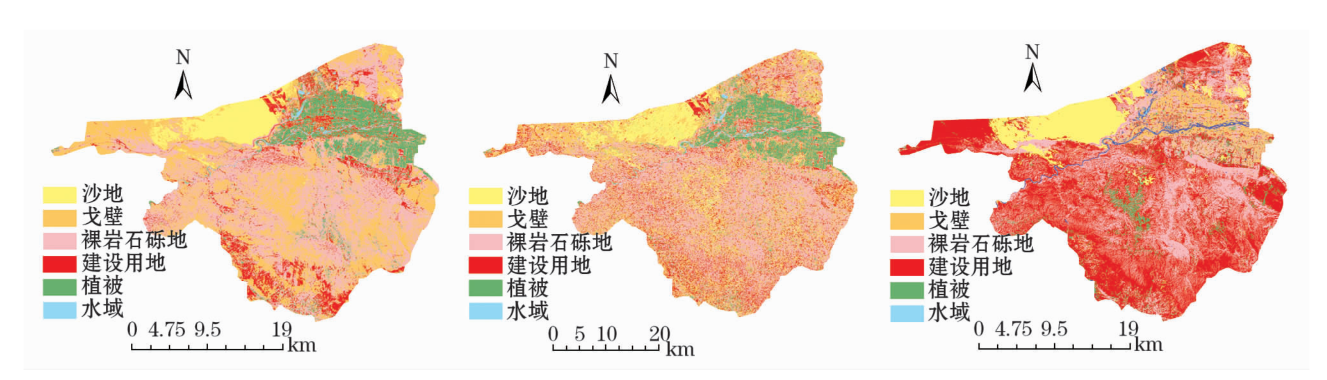


图 6　CART 决策树分类结果(a),最近邻分类结果(b)和支持向量机分类结果(c)

Fig. 6　CART decision tree classification results (a), nearest neighbor classification results (b) and support vector machine classification results (c)

表 1　三种算法的混淆矩阵精度评价

Tab. 1　Accuracy evaluation of confusion matrix of three algorithms

类别	沙地生产者精度	沙地用户精度	总体分类精度	Kappa 系数
CART 决策树算法	1.00	0.98	0.77	0.72
最近邻算法	0.66	0.98	0.69	0.63
支持向量机算法	1.00	0.86	0.49	0.38

沙地的提取方面值得借鉴推广,沙地的生产者精度分别是它们的 1.52 倍、1 倍,用户精度分别是它们的 1 倍、1.14 倍。

(2) CART 决策树采用贪心算法,虽然输入更多的提取特征对最优决策树的构建不产生过多影响,但减少不必要的特征输入可以提高提取效率和自动化程度。本试验中共输入了 22 个提取特征,最后得出 *NDBI*、*GSI* 和 *SWIR* 2(第七波段)均值是反映沙地信息的重要特征量。所筛选出来的特征为以后的沙地信息提取研究中可提供一定的参考价值。

(3) CART 决策树是一种监督分类方法,样本的质量对分类精度至关重要。由于研究条件限制,若训练样本和验证样本均为实地采集将有助于提高试验的准确性和分类结果的精确性。

参考文献(References)

[1] 王梅梅,朱志玲,吴咏梅. 宁夏中部干旱带土地沙漠化评价[J]. 中国沙漠,2013,33(2):320 - 324. [WANG Meimei,ZHU Zhiling,WU Yongmei. Assessment on the sensitivity to aeolian desertification and importance of controlling aeolian desertification in the middle arid region of Ningxia[J]. Journal of Desert Research,2013,33(2):320 - 324.]

[2] 李元科,全志杰,吕恒,等. GIS 支持下的盐池县土地沙漠化动态遥感监测与预估[J]. 干旱环境监测,1998,(4):213 - 217 + 252. [LI Yuanke,QUAN Zhijie,LV Heng,et al. Remute sensing monitoring and forecast of the desertification trends of Yanchi County by GIS[J]. Arid Environmental Moitoring,1998,(4):213 - 217,252.]

[3] 李晓琴,张振德,张佩民. 格尔木土地荒漠化遥感动态监测研究[J]. 国土资源遥感,2006,(2):61 - 63,78. [LI Xiaoqin,ZHANG Zhende,ZHANG Peimin. Remote sensing survey and monitoring of desertification in Golmud area[J]. Remote Sensing for Land & Resources,2006,(2):61 - 63,78.]

[4] 乔平林,张继贤,林宗坚. 基于神经网络的土地荒漠化信息提取方法研究[J]. 测绘学报,2004,33(1):58 - 62. [QIAO Pinglin,ZHANG Jixian,LIN Zongjian. An artificial neural network method for the information of desertification extraction[J]. Acta Geodaetica et Cartographica Sinica,2004,33(1):58 - 62.]

[5] 王志波,高志海,王琫瑜,等. 基于面向对象方法的沙化土地遥感信息提取技术研究[J]. 遥感技术与应用,2012,27(5):770 - 777. [WANG Zhibo,GAO Zhihai,WANG Bengyu,et al. The study of extracting sandy lands information from remote sensing image based on object-oriented method[J]. Remote Sensing Technology and Application,2012,27(5):770 - 777.]

[6] 冯益明,郑冬梅,智长贵,等. 面向对象的沙化土地信息提取[J]. 林业科学,2013,49(1):126 - 133. [FENG Yiming,ZHENG Dongmei,ZHI Changgui,et al. Desertification land information extraction based on object-oriented classification method[J]. Scientia Silvae Sinicae,2013,49(1):126 - 133.]

[7] 孙建伟,王超,王娜,等. 基于 CART 决策树的 ZY-3 卫星遥感数据土地利用分类监测[J]. 华中师范大学学报(自然科学版),2016,50(6):937 - 943. [SUN Jianwei,WANG Chao,WANG Na,et al. Research on land use classification monitoring though the remote sensing data of ZY-3 satellite based on CART decision tree[J]. Journal of Central China Normal University (Natural Science Edition),2016,50(6):937 - 943.]

[8] 段英杰,何政伟,王永前,等. 基于遥感数据的西藏自治区土地沙漠化监测分析研究[J]. 干旱区资源与环境,2014,28(1):55 - 61. [DUAN Yingjie,HE Zhengwei,WANG Yongqian,et al. Monitoring land desertification of Tibet Autonomous Region based on remote sensing[J]. Journal of Arid Land Resources and Environment,2014,28(1):55 - 61.]

[9] 康文平,刘树林,段翰晨. 基于 MODIS 时间序列数据的沙漠化遥感监测及沙漠化土地图谱分析——以内蒙古中西部地区为例[J]. 中国沙漠,2016,36(2):307 - 318. [KANG Wenping,LIU Shulin,DUAN Hanchen. Monitoring and spatial-temporal changes analysis of aeolian desertified lands based on MODIS data[J]. Journal of Desert Research,2016,36(2):307 - 318.]

[10] 吕利利,颉耀文,黄晓君,等. 基于 CART 决策树分类的沙漠化信息提取方法研究[J]. 遥感技术与应用,2017,32(3):499 - 506. [LYU Lili,XIE Yaowen,HUANG Xiaojun,et al. Desertification information extraction method research based on the CART decision tree classification[J]. Remote Sensing Technology and Application,2017,32(3):499 - 506.]

[11] 黄慧萍. 面向对象影像分析中的尺度问题研究[D]. 北京:中国科学院遥感应用研究所,2003. [HUANG Huiping. Scale issues in object-oriented image analysis[D]. Beijing: Institute of Remote Sensing Applications, Chinese Academy of Sciences, 2003.]

[12] 文斯. 遥感影像数据的面向对象分类与模糊逻辑分类研究[D]. 昆明:昆明理工大学,2011. [WEN Si. Remote sensing image data object-oriented classification and fuzzy logic classification research[D]. Kunming: Kunming University of Science and Technology,2011.]

[13] 张晓娟,杨英健,盖利亚,等. 基于 CART 决策树与最大似然比法的植被分类方法研究[J]. 遥感信息,2010,(2):88 - 92. [ZHANG Xiaojuan,YANG Yingjian,GAI Liya,et,al. Research on vegetation classification method based on CART decision tree algorithm and maximum likelihood ratio[J]. Remote Sensing Information,2010,(2):88 - 92.]

[14] 齐乐,岳彩荣. 基于 CART 决策树方法的遥感影像分类[J]. 林业调查规划,2011,36(2):62 - 66. [QI Le,YUE Cairong. Remote sensing image classification based on CART decision tree method[J]. Forest Inventory and Planning,2011,36(2):62 - 66.]

[15] ROUSE J W,HAAS R H,SCHELL J A,et al. Monitoring vegetation systems in the great plains with ERTS[C]//Proceeding of 3rd ERTS-1 Symposium. NASA: Goddard Space Flight Center,1974: 309 - 317.

[16] 徐涵秋. 基于谱间特征和归一化指数分析的城市建筑用地信息提取[J]. 地理研究,2005,24(2):311 - 320,324. [XU Hanqiu. Fast information extraction of urban built-up land based on the analysis of spectral signature and normalized difference index[J]. Geographical Research,2005,24(2):311 - 320,324.]

[17] 查勇,倪绍祥,杨山. 一种利用 TM 图像自动提取城镇用地信息的有效方法[J]. 遥感学报,2003,7(1):37 - 40,82. [ZHA Yong,NI Shaoxiang,YANG Shan. An effective approach to automatically extract urban land-use from TM imagery[J]. Journal of Remote Sensing,2003,7(1):37 - 40,82.]

[18] XIAO J,SHEN Y,RYUTARO T,et al. Detection of land desertification and topsoil grain size using remote sensing[J]. IEEE,2005,3(2):581 - 589.

[19] 吴宏安,蒋建军,张海龙,等. 比值居民地指数在城镇信息提取中的应用[J]. 南京师大学报(自然科学版),2006,29(3):118 - 121. [WU Hongân,JIANG Jianjun,ZHANG Hailong,et al.

Application of ratio resident area Index to retrieve urban residential areas based on Landsat TM data[J]. Journal of NanJing Normal University (Natural Science), 2006, 29(3): 118 - 121.]
[20] 苏晓玉,甘甫平,万里飞,等. 辅以波谱分析的高分辨率影像面向对象分类研究[J]. 图学学报, 2012, 3(1): 73 - 79. [SU Xiaoyu, GAN Fuping, WAN Lifei, et al. The objecct-oriented classification of high resolution images with spectrum analysis[J]. Journal of Graphics, 2012, 3(1): 73 - 79.]

Sand information extraction method based on CART decision tree

ZHANG Xi-wei[1], WANG Lei[2], WANG Xi-yuan[1,3]

(1 *School of Physics and Electronic-Electrical Engineering, Ningxia University, Yinchuan* 750021, *Ningxia, China*; 2 *Breeding Base for State Key Laboratory of Land Degradation and Ecological Restoration in Northwest China, Yinchuan* 750021, *Ningxia, China*; 3 *Ningxia Hui Autonomous Region Desert Information Intellisense Key Laboratory, Ningxia University, Yinchuan* 750021, *Ningxia, China*)

Abstract: This paper used the object-oriented method and CART decision tree method to extract the sand information with high degree of automation and comprehensive extraction features. The main research process is as follows: (1) Select the study area and preprocess the image of the study area. (2) Use multi-scale segmentation and spectral difference segmentation to obtain the object layer. (3) Select rich extraction features and training sample objects. (4) Training features and sample objects to get the CART rule tree. (5) Apply all objects to the rule tree to get the classification result. (6) Compare the Nearest neighbor and Support vector machine classification results. Finally, Compared with the current research on extracting sand information by CART decision tree. The overall classification accuracy reached 77%, which is 1.12 times of the Nearest neighbor classification result, 1.57 times of the support vector machine classification result. In addition, normalized diffevence bare index (*NDBI*), granularity size index (*GSI*) and the seventh band (*SWIR* 2) can successfully distinguish three easily mixed objects of sand, Gobi and bare rock, which are three important characteristic indexes in the process of sand extraction. The experiment has proven this method is a feasible sand extraction method for actual desertification monitoring.

Key words: object-oriented; multi-scale; spectral differences; CART decision tree; sand extraction